

Un système d'aide aux choix des modeles économetriques

Z. BOUZIDI, R. BACHA, H. KHERBACHI - Laboratoire LITEQ, Université de Béjaïa - Targa Ouzemmour, Béjaïa (06000), ALGERIE

Abstract

The contribution of artificial intelligence in econometric analysis is becoming significant. Systems with developed knowledge basis go from those solely provided for consultation where the program or an effective procedure run is carried out by the user himself, up to connecting between purely declaratives knowledge and external programs, possibly by the way of various models of knowledge representation in a given field of application.

In this project, we propose an aid system for econometric models choice: S. A. C. M. E, based on an interdisciplinary expert system approach and on the integration of knowledges, will it be qualitative or quantitative data.

First, we have elaborated a multiple regression professional software, named : mulreg, developed in a Matlab environment, by considering a tool «task» for econometric modelling. Then, we have used an expert system generating of «Turbo Prolog» with an interface program written in Turbo C language, which represent an interactive interface between. S. A. C. M. E system, mulreg professional software and the external interfaces.

It is a prototype which takes into account, simultaneously, various aspects of analysis : colinearity, autocorrelation of residues, statistic tests...

The integrated aid module, in hypertext document, makes it behave as a pedagogue.

Finally, it gives not only an idea on the model choice but also provides suggestions and decision making.

keywords: aid system, expert approach, econometric modelling...

1. Introduction

Notre travail s'inscrit dans le cadre d'un projet global initié par le Laboratoire d'Informatique et des Techniques Qualitatives «LITEQ» de l'université de Béjaïa et ayant pour intitulé «Modélisation des connaissances méthodologiques».

Il est fondé sur l'approche «système expert» interdisciplinaire et sur l'intégration des connaissances de nature qualitative et des données chiffrées.

L'objectif essentiel de celui-ci est la réalisation d'un système expert interactif, -nommé **S. A. C. M. E-**, qui permet la construction et l'utilisation des modèles économétriques. Il juge la signification globale du modèle, interprète clairement les résultats de l'analyse et élabore des prévisions.

Notre système peut guider ainsi l'utilisateur pour comprendre les évolutions des grandes variables caractéristiques d'un phénomène réel modélisable, citons: l'environnement commercial de l'entreprise (chiffre d'affaire, volume du marché, plan marketing...), l'environnement médical, secteurs publics.

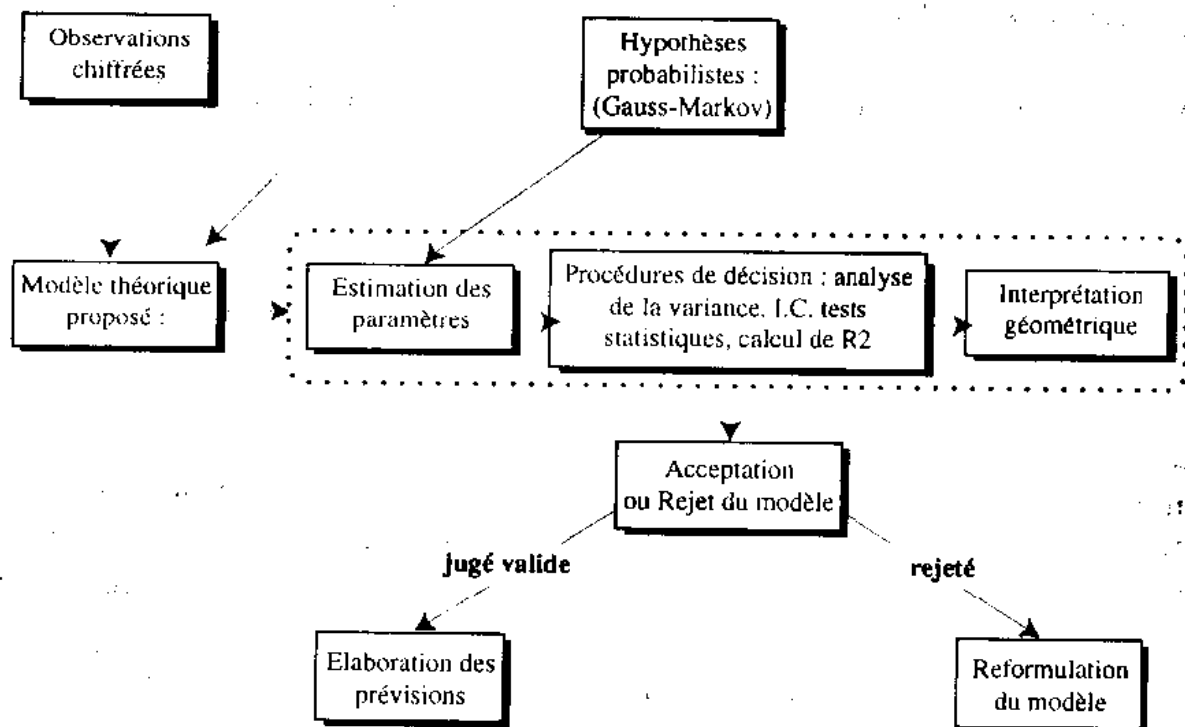
2. Eléments d'économétrie

L'économétrie se propose d'étudier par une démarche purement scientifique, les phénomènes modélisables (les questions économiques, sociales... etc).

Un modèle économétrique est simplement un modèle algébrique où intervient une erreur aléatoire. Elle synthétise l'ensemble des variables inconnues qui échappent à l'analyste et qui interviennent dans le phénomène étudié.

La forme matricielle d'un modèle linéaire est $Y = B X + \xi$ tel que :

La grandeur Y est appelée variable expliquée - endogène - ou encore une variable déterminée. Les B_i sont les paramètres inconnus à estimer. X est la matrice des variables explicatives .



Déscription des étapes de la démarche économétrique

La démarche économétrique est un processus complexe qui implique un grand nombre de connaissances et de choix. Elle comporte généralement les étapes suivantes: [GIR 94]

- Une fois le modèle économétrique spécifié par l'économetre et confronté aux observations chiffrées, on soumet cette formule mathématique aux hypothèses de Gauss-Markov qui sont : l'homoscédasticité, la non autocorrélation des résidus, la condition de linéarité et la normalité des résidus.
- Il sera ensuite estimé par la méthode appropriée (MCO, MCG, DMC) en évaluant les paramètres inconnus.
- L'étape suivante consiste à effectuer l'analyse de la variance, à étudier les régions de confiance pour les estimateurs et à établir les tests statistiques. Ces traitements se font dans le but de tester la qualité de l'ajustement linéaire entre les variables et sur la signification globale de la régression.
- L'interprétation géométrique permet de mieux visualiser parmi les variables explicatives, celles qui ont un rôle effectif de cause à effet.
- On aboutit à une décision concernant l'acceptation ou le rejet du modèle. Si le modèle est jugé valide alors il est possible d'élaborer des prévisions. Dans le cas contraire, il va falloir modifier la spécification du modèle.

3. Environnement de développement et de programmation

Avant de présenter les différents aspects à considérer pour la conception et la mise en oeuvre de notre système, nous avons pensé qu'il est préférable de décrire d'abord l'environnement de développement et de programmation de celui-ci.

Le logiciel est ainsi développé dans l'environnement Turbo PROLOG. Les raisons essentielles du choix de ce dernier sont multiples : d'abord, il possède un générateur de système expert qui nous permet de gagner un certain temps au début du développement.

Il possède un éditeur de texte intégré qui permet la gestion des menus déroulants et la création d'une interface conviviale qui répond aux exigences de base des interactions homme / machine. Il offre également la modularité et la possibilité de compiler des programmes séparément. Cependant, il reste inefficace pour les traitements numériques, particulièrement le calcul matriciel.

Cet inconvénient nous a poussé à développer un progiciel en régression multiple «MULREG» dans un autre environnement de programmation : MATLAB. Celui ci, est conçu pour ce type de calcul, tout en combinant la puissance et la rapidité d'exécution (quelques ms pour effectuer des milliers d'opérations).

Il offre également la fiabilité des résultats et une visualisation avancée dans un environnement interactif intégré.

Une autre caractéristique importante de MATLAB, est sa propriété de «système ouvert», c'est à dire la possibilité d'importer et d'exporter des données dans les environnements externes. Parmi ceux ci, nous citons Pascal, Fortran, TC.

Nous avons choisi ce dernier (le TC) pour établir le module d'interfaçage qui doit connecter le logiciel développé avec Turbo Prolog et le progiciel élaboré dans Matlab.

4. Conception

4. 1. Objectifs de SACME

L'objectif du système est de permettre la construction et l'utilisation d'un modèle économétrique.

Le système S. A. C. M. E doit interpréter clairement les résultats obtenus lors de l'analyse, porter un jugement sur le caractère significatif ou non de telle ou telle variable parmi celles retenues.

Un autre objectif de S. A. C. M. E est d'autoriser l'utilisateur à élaborer des prévisions, dans le cas où le modèle est jugé valide.

Compte tenu de la complexité du processus de la démarche économétrique, nous avons alors opté pour l'approche experte, tout en donnant l'aspect pédagogique et apprentissage au système.

L'approche experte présente plusieurs avantages : d'abord, il s'agit d'un système qui satisfait l'utilisateur en se comportant en pédagogue. Ceci grâce au module d'explication et au module d'aide élaboré dans un logiciel hypertexte. Il assure une interaction durant les étapes de résolution. Cette approche facilite également la mise à jour des connaissances évolutives et qui sont soumises souvent à des modifications.

4. 2. Modélisation des connaissances

Les connaissances de notre domaine reposent sur données chiffrées et des informations qualitatives, qui impliquent deux types de modélisation. Les connaissances purement déclaratives sont modélisées à l'aide d'un formalisme de représentation des connaissances qui sont les règles de production. L'ensemble de ces règles adoptent d'une manière générale la syntaxe suivante :

- *Ri*

- *Consulter un ou plusieurs faits (réponses de l'utilisateur, faits retournés par le progiciel)*

- *Appel à une ou plusieurs règles suivantes.*

La règle déclenchée pourra consulter en premier lieu le choix de l'utilisateur préalablement stocké en base de faits, lire un indice retourné par MULREG et appeler les règles qui suivent. Par la suite, un ensemble de faits seront sauvegardés dans le fichier «Rapport».

Quant aux connaissances quantitatives, celles-ci sont modélisées à l'aide d'un outil «tâche».

Une tâche est un outil de modélisation des connaissances, elle est définie par un nom , un ensemble d'entités d'entrées qui sont tous les objets qui doivent intervenir pour effectuer la tâche, un ensemble d'entités de sortie qui sont tous les objets résultants une fois que la tâche est exécutée et une stratégie de résolution de problèmes. [BOU 94]

< NOM > : Tâche

E : (E1, E2, ..En)

St : (St1, St2, ...Stp)

S : (S1, S2, ...Sm)

Définition d'une tâche

Le procédé de la modélisation de la démarche économétrique consiste en fait à construire un arbre hiérarchique dont la racine est la tâche principale. Celle-ci génère des branches «sous-tâches». Les tâches se décomposent en sous tâches et le processus de décomposition se poursuit jusqu'à arriver aux feuilles qui sont des tâches élémentaires et indécomposables.

Hormis les feuilles, les autres éléments de l'arbre sont des tâches complexes. La caractéristique d'une tâche complexe est qu'elle est décomposable, en générant des sous tâches qui peuvent être séquentielles, alternatives ou encore itératives.

Nous avons par la suite intégré ces connaissances modélisées dans un progiciel qu'on a nommé MULREG que nous avons réalisé dans l'environnement Matlab.

Ce progiciel doit contrôler l'exécution de ces tâches soit en faisant intervenir le choix de l'utilisateur ou à des calculs et traitements.

5. Réalisation

5. 1. Composants de SACME

Un Système Expert est un logiciel qui dans le cadre d'un domaine d'application déterminé, résout les problèmes d'un utilisateur d'une manière proche de celle d'un expert. Notre SE est constitué - comme tout autre SE - d'une base de connaissances rassemblant les faits connus et les règles d'inférence et d'un moteur d'inférence qui exploite les faits connus et ces règles afin d'arriver à la solution.

Notre moteur d'inférence est celui de Turbo Prolog, qui fonctionne en logique d'ordre «un» et raisonne en chaînage arrière.

La base de faits se remplit à l'aide des réponses données par l'utilisateur aux questions posées par le système ainsi que les déductions établies par ce dernier.

Elle se construit aussi à partir des résultats quantitatifs retournés par le progiciel MULREG développé dans Matlab.

Quant à la base de règles, nous avons adopté deux paquets de règles, l'un concerne les modèles à équations uniques et l'autre concerne les modèles à équations simultanées.

Pour accéder à l'un des deux paquets de règles, il est nécessaire d'utiliser une métarègle qui va distinguer le type du modèle saisi et sélectionner l'un des deux paquets.

Sa syntaxe ne diffère pas de celle d'une règle normale :

SI ($n=1$) ALORS (Sélectionnez le paquet de règles concernant les modèles à équations uniques) SINON (Sélectionnez le paquet de règles concernant les modèles à équations simultanées)

n étant le nombre de variable endogènes qui est le nombre de relations structurelles du modèle saisi

Syntaxe de la métarègle

Notre système est conçu également autour d'un certain nombre de modules et interfaces.

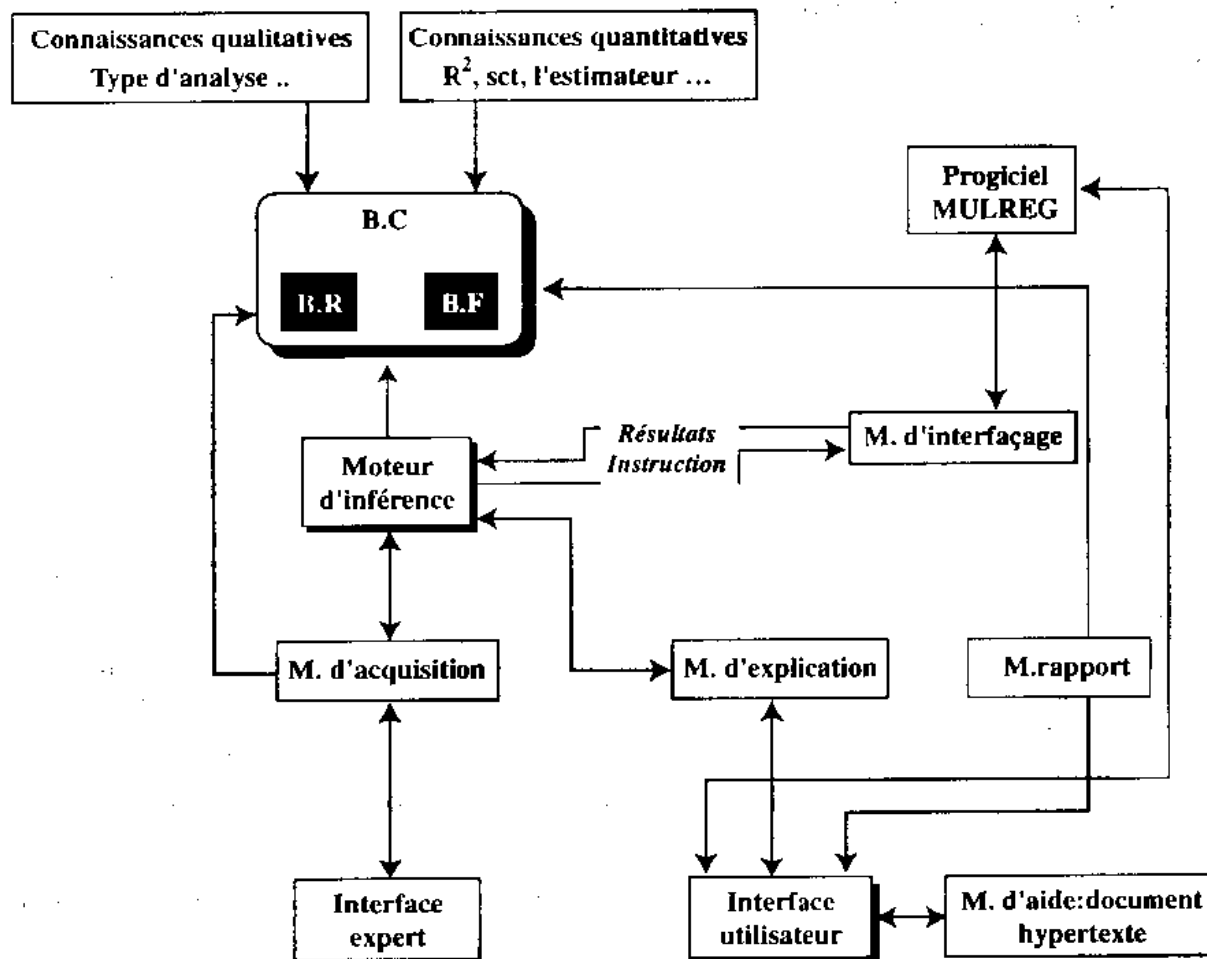
On distingue les interfaces utilisateur et expert, l'interface avec le logiciel qui permet une translation fiable des données entre le système et le logiciel.

L'interface expert offre la possibilité à l'expert humain d'accéder à la base de connaissances pour une consultation ou une mise à jour des règles ou des faits. Ceci est réalisé à l'aide du module d'Acquisition qui doit contrôler si les règles ou les faits introduits ou modifiés sont conformes à la syntaxe prédéfinie et ne contredisent pas celles qui existent déjà.

La tâche principale du module d'Explication est de satisfaire l'utilisateur en répondant aux questions qui apparaissent ambiguës aux yeux de l'utilisateur.

Le module «Rapport» est un recueil de toutes les conclusions du modèle saisi, qui sont regroupées dans un fichier qu'on peut visualiser sur écran ou l'imprimer sur papier.

Nous avons tenu compte de l'aspect pédagogique du système en intégrant le module hypertexte d'aide. Ce dernier permet non seulement de fournir à l'utilisateur des explications à propos du fonctionnement général du logiciel, mais aussi permet d'acquérir des notions vagues sur l'économétrie jusqu'aux mots clés de cette discipline



Architecture de S.A.C.M.E

5. 2. Fonctionnement

Voici une interprétation brève de son fonctionnement.

L'utilisateur saisit le modèle économétrique, à travers l'interface utilisateur. Le moteur d'inférence détermine la nature de celui-ci et affiche la méthode envisageable pour l'analyse de ce type de modèle, en faisant intervenir le choix de l'utilisateur.

Ensuite, il fournit ces instructions au progiciel MULREG, à travers le module d'interfaçage qui transmettra celles-ci dans l'environnement du progiciel.

Durant l'analyse du modèle, le progiciel demande à l'utilisateur de lui compléter certaines informations supplémentaires, et lui affiche les résultats numériques obtenus.

En fin de traitement, les résultats seront transmis à SACME via le module d'interfaçage.

Le système expert entame alors la phase d'interprétation à partir des informations qualitatives et des données quantitatives, ceci à l'aide d'un mécanisme d'exploitation de ces informations.

En effet, le moteur d'inférence déclenche la métarègle qui permet de distinguer le type du modèle (modèle à équations uniques / modèle à équations simultanées) et sélectionner le paquet de règles qui convient, et ce en consultant les connaissances fournies par l'utilisateur lors de la saisie et celles retournées par le progiciel et préalablement stockées en base de faits.

En fin de traitement, le système communique son diagnostic à l'utilisateur à travers le module Rapport.

Il arrive que l'utilisateur se pose certaines questions vagues ou particulières, il recevra d'amples explications grâce au module d'Explication et au module d'Aide.

Voici un exemple tiré du module d'explication, il s'agit d'une question posée par le système à l'utilisateur.

6. Conclusion

L'approche experte a toujours été l'outil favori pour la résolution des problèmes aussi complexes soient-ils.

Elle intervient en effet dans bon nombre d'activités intellectuelles, parmi lesquelles, on cite l'économétrie, qui a été l'objet de notre étude.

La difficulté principale rencontrée, pendant l'élaboration de S. A. C. M. E est le problème d'occupation de la mémoire vive, dû à l'utilisation conjointe de deux logiciels S. A. C. M. E et mulreg.

Ce problème est résolu par la programmation modulaire que Turbo Prolog nous offre et par la manière dont on a modélisé la démarche économétrique.

Nous avons également trouvé des difficultés pour la connexion du S. E «S. A. C. M. E» avec le progiciel MULREG, il a fallu alors établir un programme d'interface écrit dans le langage Turbo C et qui assure une communication permanente et fiable.

Nous pensons avoir réalisé un prototype de système expert consacré à l'analyse des modèles économétriques. Il tient compte simultanément de plusieurs aspects de l'analyse (colinéarité, l'autocorrélation des résidus) mais celui-ci demeure ouvert à de nombreuses extensions, dans le but de le rendre aussi ouvert que possible.

Il est ainsi souhaitable de stocker la base de données d'informations qualitatives en employant une norme des faits plus puissants.

Nous souhaitons aussi introduire les nouvelles méthodes de prévisions, à moyen et à long terme pour les modèles à équations simultanées.

Enfin, nous aimerions reprendre ce travail, cette fois, en optant pour un environnement de programmation sous windows.

Il est vrai que certains systèmes ont rencontré plus de succès que d'autres, mais aucun d'entre eux ne peut cependant être considéré comme totalement satisfaisant, car comme dit *Norman Meyrowitz* «*Il reste un joli bout de chemin à faire avant d'avoir un système parfait.* »

7. Références bibliographiques

- [BOU 94] Z. Bouzidi, H. Kherbachi & A. Hocine
Intégration des connaissances en analyse statistique
LICIAP, institut d'informatique (université de Pau et des Pays de l'Adour)
LITEQ, institut d'économie (centre universitaire de Béjaïa «Algérie»)
- [BRU& 88] H. de la Bruslerie, J. Rigal
Informatique : Aide à la décision de gestion
Edition Vuibert 1988
- [GIR 94] René Giraud et Nicole Chaix
Econométrie
Edition Economie 1994
- [HOC 93] A. Hocine
Développement des SYSTÈME EXPERT dans l'environnement PROLOG
Laboratoire d'Informatique LICIAP
Université de Pau et des pays de l'Adour 1993
- [LAB 85] Christian Labrousse
Introduction à l'économétrie
Edition Dunod 1985
- [LAF& 90] Annie Lafort, Guy Melard et T. M. Pasteels
Vers un système de prévision et de statistiques économiques
Cahier du C. E. R. O Volume N° 32 Institut de Statistiques
Université Libre de Bruxelles - Belgique 1990
- [ROB 91] Claudine Robert
*Modèles statistiques pour l'intelligence artificielle:
l'exemple de diagnostic médical.* Edition Masson 1991.